GENETIC SEQUENCES, DIAGNOSTIC AND/OR QUANTIFICATION METHODS AND DEVICES FOR THE IDENTIFICATION OF *STAPHYLOCOCCI* STRAINS

Field of the invention

The present invention refers to new genetic sequences, diagnostic and/or quantification methods and devices using said sequences for the identification of various types of *Staphylococci* strains as well as the therapeutical aspects of said genetic sequences.

Background of the invention

Increasing incidence of nosocomial infections by multiresistant bacteria (even to antibiotics like vancomycin) is a world-wide concern. Methicillin-resistant coagulase-negative *Staphylococci* (MR-CNS) and *S. aureus* (MRSA) express a high level cross-resistance to all ß-lactam antibiotics (Ryffel et al. (1990), Refsahl et al. (1992)). They have an additional low-affinity penicillin-building protein, PBP2a (PBP2'), encoded by the *mecA* gene. The *mecA* determinant is found in all multiresistant staphylococcal species (Chackbart et al. (1989), Suzuki et al. (1992), Vannuffel et al. (1995)) and is highly conserved among the different species (Ryffel et al. (1990)).

Several other chromosomal sites, in which transposon inactivation reduces the level of ß-lactam resistance, have been identified in *S. aureus* (SA) (Hiramatsu (1992), Berger-Bächi et al. (1992), de Lancastre et al. (1994)). The appropriate functioning of these regulator genes rather than the quantity of PBP2a determines the minimal inhibitory concentration value and homogeneous expression of resistance of staphylococcal isolates (Ryffel et al. (1994), de Lancastre et al. (1994)).

The *femA-femB* operon, initially identified in *S. aureus*, is one of those genetic factors essential for methicillin resistance (Berger-Bächi et al. (1989)). It is involved in the formation of the characteristic pentaglycine side chain of the SA peptidoglycan (Stranden et al. (1997)). Unlike other regulatory genes, *femA* was shown to retain a strong conservation over time in clinical isolates of MRSA, hence confirming its key role in cell wall metabolism and methicillin resistance (Hurlimann-Dalel et al. (1992)). In contrast to *mecA*, *femA-femB* is present both in the genome of resistant and susceptible SA strains (Unal et al. (1992), Vannuffel et al. (1995)).

Often, identification of the *Staphylococci* is limited to a rapid screening test for *S. aureus*, and non-*S. aureus* isolates are simply reported as coagulase-negative *Staphylococci*. In fact, these bacteria isolates include a variety of species and many different strains (Kleeman et al. (1993)). There is little epidemiological information related to the acquisition and spread of these organisms. This is potentially due to the lack of an easy and accurate way to identify species and to provide clinically timely informations.

Several molecular assays designed for detecting *femA* in SA failed to amplify an homologous sequence in coagulase-negative *Staphylococci* (Kizaki et al. (1994), Vannuffel et al. (1995)). Nevertheless, low-stringency heterologous hybridisation analysis suggested the presence of such a structurally related gene in *S. epidermidis* (SE) (Unal et al. (1992)).

These data were followed by complete identification and sequence analysis of the *femA* and *femB* open reading frames in *S. epidermidis* (Alborn et al. (1996)). Intra- and interspecies relatedness of these genes and conservation of genomic organisation are therefore consistent with gene duplication of one of these genes in an ancestral organism and the possibility of *femA* phylogenetic conservation in all staphylococcal species (Alborn et al. (1996)).

The complete genetic sequence of the *femA* gene de *S. epidermidis*, the protein encoded by the *femA* gene (*FemA*) and vectors and micro-organisms comprising genes encoding the *FemA* protein are described in the US patent 5,587,307.

Aims of the invention

The present invention aims to provide new genetic sequences, methods and devices for the improvement of the identification and/or the quantification of various types of *Staphylococci* strains through their *femA*-like determinants, which allow by a rapid screening their epidemiological study.

Another aim of the invention is to identify similar genetic sequences which may exist in known or not

known *Staphylococci* species or other gram-positive bacterial strains.

A last aim of the present invention is to provide new sequences encoding *femA* proteins of *Staphylococci* species, their *femA* proteins, vector(s) comprising said nucleotide sequences and cell (s) transformed by said vector(s) for possible therapeutical applications.

Summary of the invention

The Inventors have identified new DNA and amino acid sequences from new strains of *Staphylococcus hominis*, *Staphylococcus saprophyticus* and *Staphylococcus haemolyticus*. Said new nucleotide sequences allow an alignment of these new sequences with the *femA* gene from *Staphylococci* previously described (*S. aureus*, *S. epidermidis* and *S. saprophyticus*). By the alignment of more than 2 sequences, preferably more than 4 sequences, the Inventors have identified for the first time a consensus *femA* sequence useful for molecular genotyping of different *Staphylococci* species which was not possible previously, when only few *femA* sequences of *Staphylococci* strains were known.

Therefore, a first aspect of the present invention is related to the "consensus" nucleotide sequence as represented in the enclosed Figure 3. With said "consensus" nucleotide sequence, the Inventors were able to provide oligonucleotides (such as primers or probes) which can be used for the genetic amplification, the identification and/or quantification of various *femA* sequences which are specific of known or unknown *Staphylococci* species.

The *femA* sequence is known to be involved with the biosynthesis of glycin-containing cross-bridges of the peptidoglycan and the peptidoglycan organisation is also known to be well conserved among various *Staphylococci* species and possibly among other gram-positive bacteria.

Therefore, it is also possible to use the new "consensus" *femA* sequence and said new oligonucleotides extrapolated from the alignment of the sequences presented in Figure 3, for the molecular genotyping of other *Staphylococci* species and possibly other gram-positive bacteria. It is also known that the *femA* sequence is similar to the femB sequence. Therefore, these oligonucleotides could also be used for the molecular genotyping of femB genes of different *Staphylococci* species or other gram-positive bacteria.

Another aspect of the present invention concerns the possible therapeutical uses of new *femA* nucleotide sequences isolated from the strains *S. hominis*, *S. saprophyticus*, *S. haemolyticus*, *S. lugdunensis*, *S. xylosus*, *S. capitis*, *S. schleiferi* and *S. sciuri* having a nucleotide or amino acid sequence which presents more than 85%, preferably more than 90% homology or 100% homology with the genetic sequences presented in the Figures 6 to 13, their complementary strand and functional variants thereof. Functional variants of said amino acid sequences are peptides which contain one or more modifications to the primary amino acids sequence and retain the activity of the complete and wild type *femA* molecule. Variants of the peptide are obtained by nucleotidic sequences which differ from the above-identified described sequences by a degeneration of their genetic code or are sequences which hybridise with said sequences or their complementary

strand, preferably under stringent conditions such as the ones described in the document Sambrook et al., §§ 9.47-9.51 in *Molecular Cloning : A Laboratory Manual*, Cold Spring Harbor, Laboratory Press, Cold Spring Harbor, New York (1989).

A further aspect of the present invention concerns the recombinant vector (i.e. constructions into which the sequence of the invention may be inserted for transport in different genetic environments and for expression in a host cell, such as a phagemide, a virus, a plasmid, a cationic vesicle, a liposome, etc.) comprising said nucleotide sequences and their complementary strands, or the corresponding RNA sequences, possibly linked to one or more regulatory sequences or markers (resistance to antibiotics, enzyme coding sequences, ...) active into a cell.

Similarly, the nucleic acid sequence according to the invention may be obtained by synthetic methodology well known by the person skilled in the art, such as the one described by Brown et al. ("Method of Enzymology", Acad. Press, New-York, No. 68 pp. 109-151 (1979)) or by conventional DNA synthesising apparatus such as the applied biosystem model 380A or 380B DNA synthesiser.

Other aspects of the present invention concern the recombinant host (prokaryotic) cell transformed by said vector and the purified (possibly recombinant) proteins or peptides encoded by said nucleic acid sequences, possibly linked to a carrier molecule such as BSA and obtained by said cells. Said recombinant proteins or peptides could be obtained by genetic engineering or could be obtained by synthesis (see US patent 5,587,307

incorporated herein by reference) and may comprise residues enhancing their stability (resistance to hydrolysis by proteases, etc.) such as the one described by Nachman et al. (*Regul. Pept. Vol.* 57, pp. 359-370 (1995)).

A preferred vector for expression in a E. coli host cell is derived from the E. coli plasmid pET-11A available from Novagen Inc. (Catalogue No. 69436-A). The transformation technique used with the above-identified vector has been described in the US Patent 5587307.

A further aspect of the present invention concerns the inhibitor (used to possibly treat (with addition of antibiotics) antibiotics resistance bacteria) directed against said proteins, peptides or nucleic acid molecules. Advantageously, said inhibitor is a antibody, preferably a monoclonal antibody, or an antisense nucleotide molecule, such as a ribozyme, which could be present in a vector in order to block the expression of said *femA* nucleotide sequences.

A last aspect of the present invention concerns the pharmaceutical composition, preferably a vaccine, against *Staphylococci* infections in an animal, including a human, comprising a pharmaceutically acceptable carrier and a sufficient amount of an active compound selected from the group consisting of said nucleic acid molecules, vectors, recombinant host cells transformed by said vector(s), inhibitors (directed against said proteins, peptides or nucleic acid molecules) and a mixture thereof.

Another aspect of the present invention concerns oligonucleotides which are (DNA) sequences having between 15 and 350 base pairs, preferably between 17 and 250 base pairs (such as primers or probes) obtained from the consensus sequence of Figure 3 or its complementary

strand. Preferably, said oligonucleotides are primers having between 15 and 45 base pairs, more preferably between 17 and 25 base pairs.

According to a first embodiment of the present invention, said oligonucleotide is a primer having between 15 and 45 base pairs, which presents more than 60%, advantageously more than 70%, preferably more than 80%, more specifically more than 90% homology with (fragments of) the "consensus" *femA* nucleotide sequence (CNS) identified in the Figure 3.

Therefore, the oligonucleotides according to the invention are new sequences or preferred fragments of known sequences of *S. aureus*, *S. epidermidis* or *S. simulans* but not the complete wild type known *femA* nucleotide sequence.

Preferably, the oligonucleotide according to the invention is selected from the group consisting of the following nucleotide sequences :

- ANAATGAANTTTACNAATTTNACNGCNANAGANTT
  and more particularly *femS1* TAATGAAGTTTACAAAATTT or *femS2* TAATGAAGTTTACNAAATTT
- ATGNCNNANAGNCATTTNACNCANA
  and more particularly *femU1* ("universal" sequence sense of the multiplex PCR): TGCCATATAGTCATTTACGC
- TAGTNGGNATNAANAANAANNATAANGANGTNATTGC
- GTNCCNGTNATGAAANTNTTNAANTANTTTTATTC
- AATGCNGGNNANGATTGG
- GNAANNGNAANACNAAAAAAGTNNANAANAATGGNGTNAAAGT
  and more particularly *fsq1S* (et *1AS*) : AAAAAGTTCAAAAAATGG and *fsq2S* (and *2AS*) : AAAAAGTACAAAAAATGG
- AAGANGANNTNCCNATNTTNNGNTCATTNATGGANGATAC

- TATATNNANTTTGATGANTA
- AANGANATNGANAAANGNCCNGANAANAAAAA

and more particularly *fsq3S* (and *3AS*) : AAAGATATTGAAAAACGA, *fsq4S* (and *4AS*) : AAAGATATTGAAAAGAGACC, *fsq5S* (and *5AS*) : AAAGATATCGAGAAAGAC and *fsq6S* (and *6AS*) : AAAGACATCGACAAGCGT.

- ANCATGGNAANGAATTACCNAT

and more particularly *fem1* (primer for the production of a probe and of marked amplicons for reverse hybridisation experiment) : GAACATGGTAATGAATTAC

- AATCCNTNTGAAGTNGTNTANTANGCNGGTGG
- AGNTATGCNNTNCAATGGNNNATGATTAANTATGC
- TTTANNGANGANGCNGAAGATGNNGGNGTNNTNAANTTNAAAAA

and more particularly *fem3bio* (primer for the production of a probe and of marked amplicons for reverse hybridisation experiment) : TTTACTGAAGATGCTGAAGA

- GTTGGNGANTTNNTNAAACC

and more particularly *fem2* (primer for the production of a probe and of marked amplicons for reverse hybridisation experiment) : GTTGGTGACTTTATTAAACC

- ATGAAATTTACAGAGTTAA (= *femAS1*)

Said primer(s) will be designated hereafter as "universal primer(s)".

A further aspect of the present invention concerns the oligonucleotide being either a primer or a probe as above-described, having between 15 and 350 base pairs, preferably between 17 and 250 base pairs, or a primer having between 15 and 45 base pairs, more preferably between 17 and 25 base pairs, which will be designated

hereafter as "specific primer(s)", having a nucleotide sequence which presents less than 50%, advantageously less than 40%, preferably less than 30%, more specifically less than 20% homology with (fragments of) the "consensus" *femA* nucleotide sequence (CNS) identified in the Figure 3 and with another *femA nucleotide* sequence specific for other *Staphylococci* strains.

Advantageously, said "specific primer" is selected from the group consisting of the following nucleotide sequences :

- ACAGCAGATGACATCATT
- TAATGAAAGAAATGTGCTTA
- ACACAACTTCAATTAGAAC
- AGTATTAGCAAATGCGG
- ATGCATATTTTCCGTAA
- CAGCAGATGACATCATT
- CATCTAAAGATATATTAAATGGA
- AGTATTAGCAAATGCGGGTCAC
- CAACACAACTTCAATTAGAA

The oligonucleotides according to the invention are selected according to their physiochemical properties in order to avoid cross-hybridisation between themselves. Said primers are not complementary to each other and they contain a similar percentage of bases GC.

Said oligonucleotides are used in an identification and/or quantification method of one or more *Staphylococcus* species and possibly other gram-positive bacteria.

Therefore, another aspect of the present invention is related to an identification and/or

quantification method of a *Staphylococci* species which may present resistance to one or more antibiotic(s), and is possibly combined with a method for the identification of a resistance to antibiotics, especially β-lactam antibiotics, (for instance through the identification of a variant of the *mecA* gene as described by Vannuffel et al. (1998)).

The method for the detection, the identification and/or the quantification of a bacteria, preferably a staphylococcal species, comprises the steps of :

- obtaining a nucleotide sequence from said bacteria present in a sample, preferably a biological body sample obtained from a patient such as blood, serum, dialyse liquid or cerebrospinal liquid, or from any other bacteriological growth medium,
- possibly purifying said nucleotide sequence from possible contaminants,
- possibly amplifying by known genetic amplification techniques said nucleotide sequence with one or more universal oligonucleotide(s) (universal primer(s)) according to the invention, and
- identifying the specific gram-positive bacteria species, preferably the specific *Staphylocossi* species :
  - by a comparative measure of the length of the (possibly amplified) nucleotide sequence or
  - by reverse hybridisation of the (possibly amplified) nucleotide sequence with one or more specific oligonucleotide(s) (specific probe(s) or primer(s)) according to the invention which are specific of said bacteria, said oligonucleotide(s) being preferably immobilised on a solid support.

The comparative measure of the length of a possibly amplified nucleotide sequences can be performed by the analysis of their migration (compared with a known ladder) upon an electrophoresis gel.

Preferably, the genetic amplification technique is selected from the group consisting of PCR (US patent 4,965,188), LCR (Landgren et al., *Sciences*, 241, pp. 1077-1080 (1988)), NASBA (Kievits et al., *J. Virol. Methods*, 35, pp. 273-286 (1991)), CPR (patent WO95/14106) or ICR.

The specific detection of the possibly amplified nucleotide sequences can be obtained by the person skilled in the art by using known specific gel electrophoresis techniques, in situ hybridisation, hybridisation on solid support, in solution, on dot blot, by Northern blot or Southern blot hybridisation, etc.

Advantageously, the probes which are specific of the bacteria are immobilised on a solid support according to the method described in the international patent application WO98/11253 incorporated herein by reference.

Said specific oligonucleotides (probes or "elongated" primers) have a length comprised between 50 and 350 base pairs, preferably between 120 and 250 base pairs, and are fixed to the solid support by a terminal 5' phosphate upon an amine function of the solid support by carbodiimide reaction (as described in the document WO98/11253 incorporated herein by reference).

The solid support can be selected from the group consisting of cellulose or nylon filters, plastic supports such as 96-well microtiter plates, microbeads,

preferably magnetic microbeads, or any other support suitable for the fixation of a nucleotide sequence.

The method according to the invention can be advantageously combined with another specific detection step of a possible resistance to antibiotics, especially β-lactam antibiotics (for instance through the identification by the above-described technique of variants of the *mecA* gene as described by Vannuffel et al. (1998)).

The present invention concerns also a diagnostic and/or quantification device or kit for the identification and/or the quantification of a *Staphylococcus* species or other gram-positive bacteria, comprising the oligonucleotides according to the invention and possibly all the media necessary for the identification of a (possibly amplified) nucleotide sequence of said bacteria through any one of the above-described methods.

Advantageously, the method and device according to the invention are adapted for the quantification of said *Staphylococci* strains by the use of a "internal or external standard sequence", preferably the one described in the patent application WO98/11253 incorporated herein by reference.

Therefore, according to a first embodiment of the present invention, the nucleic acid sequence from a *Staphylococcus* species, for instance *Staphylococcus aureus*, is amplified by a "universal primer" and by a "specific primer" which is specific for *S. aureus*. The identification of *S. aureus* will be obtained upon an agarose electrophoresis gel wherein the amplified nucleotide sequence (shorter than the amplified nucleotide sequence of another *Staphylococci* species such as *S. epidermidis*) and identified by the use of a comparative ladder.

According to another embodiment of the present invention, a *Staphylococcus* species (such as *S. aureus*) is identified by reverse hybridisation of the amplified nucleotide sequence with a probe which is specific of said bacteria and which is immobilised on a solid support such as filter.

The present invention will be described in details in the following non-limiting examples, in reference to the Figures described hereafter.

Short description of the drawings

The Figure 1 represents 5 partially overlapping fragments of the *femA* genes from *S. hominis*, *S. saprophyticus* and *S. haemolyticus* obtained by PCR amplification.

The Figure 2 represents the alignment of the nucleotide sequences of *femA* genes from *S. hominis*, *S. saprophyticus*, *S. aureus*, *S. epidermidis* and *S. haemolyticus*.

The Figure 3 represents the consensus sequence according to the invention.

The Figure 4 represents the result of differential diagnosis between different strains of *Staphylococci* by reverse hybridisation.

The Figure 5 represents amplification of CNS species under universal conditions.

Figures 6 to 13 represent the complete *femA* wild type genetic sequence of the strains *S. hominis*, *S. saprophyticus*, *S. haemolyticus*, *S. lugdunensis*, *S. xylosus*, *S. capitis*, *S. schleiferi* and *S. sciuri*.

Examples

Example 1 : Sequencing strategy

Fragments of the *femA* genes from S. *hominis* and S. *saprophyticus* have been obtained by PCR amplification, in low stringency annealing conditions. Primers used for amplification are matching the potentially conserved regions and have been designed according to sequences homologies between *S. aureus*, *S. sapropyticus* and *S. epidermidis femA* nucleotide sequences. For both *S. hominis* and *S. saprophyticus* species, 5 partially overlapping fragments have been synthesised allowing the sequencing of the entire *femA* genes (Fig. 1).

Example 2 : Identification of a consensus sequence

Alignment of the nucleotide sequences of *femA* genes from *S. hominis* and *S. saprophyticus* as well as with *femA* genes sequenced to date from *S. aureus* (GenBank accession number M23918), *S. epidermidis* (GenBank accession number U23713) and *S. haemolyticus* is presented in Fig. 3 and has allowed to propose a "consensus" *femA* nucleotide sequence (CNS) whose genomic organisation displays highly conserved regions flanked by variable ones. On this basis, interspecies phylogenetic variations could be exploited to design genotyping strategies for species-specific identification of *Staphylococci*. The "consensus" sequence is therefore a powerful molecular tool for specific diagnostic of staphylococcal infections.

Example 3 : Sequencing of other staphylococcal *femA* genes

The consensus sequence can be exploited for designing universal primers allowing the production, under permissive annealing conditions, of overlapping PCR

products whose sequencing will identify the entire *femA* sequence.

Example 4 : Differential diagnosis between *S. aureus, S. epidermidis, S. hominis* and *S. saprophyticus* by reverse hybridisation

The Inventors have set up a reverse hybridisation assay for rapid and combined identification of the most clinically relevant *Staphylococci* species, and their *mecA* status. Two sets of primers, chosen in a conserved domain of the consensus sequence (*bioU1-bioU3* and *fem1-fem3bio*), amplifying a 286 and bio-220 bp fragments, respectively) were synthesised. Species-specificity of *femA* amplicons was insured by the genomic variability between the conserved regions. *FemA* probes were immobilised on nylon strips. Hybridisation was performed with biotinylated *femA* PCR fragments from the strain of interest. The strategy was first assessed with ATCC strains (*S. aureus, S. epidermidis, S. hominis* and *S. saprophyticus*) (Fig. 4). Specificity was identified by standard methods. Accuracy was 100% for species identification.

Example 5 : Differential diagnosis between staphylococcal species

This assay is able to identify any staphylococcal species if following requirements are fulfilled :

- primers *fem1, fem2* and *fem3bio* are universal for *Staphylococci;*
- there is a wide enough phylogenetic variation between any CNS species to promote a specific hybridisation.

The first requirement is fulfilled for, i.e., *S. haemolyticus*, *S. capitis*, *S. cohnii*, *S. xylosus*, *S. simulans*, *S. lugdunensis*, *S. schleiferi* and *S. warneri* strains (Fig. 5).

Example 6 : Multiplex amplification of *femA* and *mecA* genetic determinants for a molecular diagnosis of a specific staphylococcal infection

A total of 48 patients treated in 4 contiguous intensive cares units were included in the study. Endotracheal aspirates (ETA) were collected from the patients and submitted to the multiplex PCR analysis according to the technique described by Vannuffel et al. (1995). Clinical specimens were homogenised in 5 ml of TE buffer (20 mM TRIS HCl, pH 8.0, 10 mM EDTA) containing 2% (w/v) SDS.

The homogenate (1.5 ml) was then centrifuged for 5 minutes at 7500 xg. The cellular pellet was washed once with TE buffer lysed in the presence of 1% (v/v) Triton X-100 and 50 µg of lysostaphin (Sigma) and incubated for 15 minutes at 37 °C. Lysis was completed by adding 100 µg of proteinase K (Boehringer). The lysate was incubated for another 5 minutes at 55 °C and 5 minutes at 95 °C, and centrifuged at 4000 xg for 5 minutes.

In order to purify bacterial DNA, 200 µl of supernatant were then filtered on a Macherey-Nagel Nucleospin C+T® column and eluted with 200 µl sterile H_2O. Two different amounts of DNA suspension (2 µl and 200 µl) were submitted to multiplex PCR amplification with the primers 5'-TGGCTATCGTGTCACAATCG-3' and 5'-

CTGGAACTTGTTGAGCAGAG-3' for *mecA* and the above-described primers for *femA*, yielding different fragments.

femA and *mecA* signals were found in specimens containing either susceptible *S. aureus* (n = 10) and methycillin-resistant coagulase-negative *Staphylococci* (n = 6) respectively. On the other hand, no signal was obtained from ETA gram-negative bacteria (n = 5) as well as MS-CNS (n = 6) and from 5 ETA containing normal pharyngeal flora.

This multiplex PCR strategy for detecting *Staphylococci* in ETA was completed in less than 6 hours either on the day of the samples' collection. This is an advantage with respect to the time required to conventional identification and susceptibility tests (48 to 72 hours).

Example 7 : Amplification, cloning and sequencing of other *femA* genes

Two primers were selected among the conserved parts of the consensus sequence for the amplification of the *femA* gene.

These primers are *femS1*, *femS2* and *femAS1* (anti-sense primer). ADN from strains of *Staphylococcus hominis*, *saprophyticus*, *haemolyticus*, *lugdunensis*, *schleiferi*, *sciuri*, *xylosus*, *simulans*, *capitis*, *gallinarum*, *cohnii* and *warneri* were amplified from said primers and amplification fragments were cloned in the vector pCR®-XLTOPO and introduced by electroporation in *E. coli* cells TOP10 (TOPO XL PCR Cloning Kit®, Invitrogen, Carlsbad, CA).

Amplified fragments of strain *S. lugdunensis*, *schleiferi*, *sciuri*, *xylosus*, and *capitis* were sequenced by Taq Dye Deoxy Terminator Cycle® sequencing on a ABI 277 DNA

sequencer® (PE Applied Biosystems, Foster City, CA) by the following primers :

femS1 or *femS2* or *femAS1*

fsq1S and *fsq1AS*

fsq2S and *fsq2AS*

fsq3S and *fsq3AS*

fsq4S and *fsq4AS*

fsq5S and *fsq5AS*

fsq6S and *fsq6AS*